Lane change behavior with a dual-task paradigm: individual differences and cognitive processes.

Abstract

Driving a car is highly complex and therefore causes driving error to be a high causal factor in road accidents. Because it requires constant management of cognitive processes like perception, decision-making, and execution of motor responses. It is therefore important to study driving behavior to ensure road safety. The paper of van der Heiden, Janssen, Donker and Merkx (2019) studied lane changing while performing a secondary task. However, the effects of primarily cognitively loading tasks on driving performance are not well understood yet. Therefore this paper extends their study with novel analyses. Firstly, this paper investigates which subscale of the TLX questionnaire is responsible for the decrease in drive performance when there is a higher level of cognitive distraction. Secondly, it is investigated if mental workload is a predictor of reaction time and distance. The third research question answers the question which underlying cognitive processes account for the different aspects of reaction time. The results showed that no specific subscale of the TLX is responsible for decrease in drive performance and there was no relationship found between reaction time, distance and mental workload. The third analysis showed drivers tend to shift from focus from the near point to the far point when they start driving faster. Furthermore drivers steer less abruptly when they drive faster with a focus on the far point. The last finding is about the delay period (the period in which the visual events are processed before action is initiated). This seems to be very important for this component T1. This information can be used to design safer roadways and enhance safety systems in cars by incorporating these cognitive processes.

Master: Applied Cognitive Psychology

Sietske Bootsma (5523230)

Supervisor: Dr. L van Maanen

1. Introduction

Understanding driving behavior is extremely important since driver distraction has become a significant source of injuries and accidents on the roadway (Strayer, Turrill, Cooper, Coleman, Medeiros-Ward & Biondi, 2015; Papantoniou, Yannis & Christofa, 2019). Drivers need to obtain a large amount of information such as road conditions, control vehicle position and speed. Managing these requirements involve information perception, decision-making, and execution of tasks. Due to the dynamic interleaving and execution of all these tasks, this is highly complex (Chong, Mirchi, Silva & Strybel, 2014; Chen, Xue, & Jiang, 2019). This complexity causes driving error to be a high causal factor in road accidents and therefore it is important to study driving behavior to ensure road safety (Dingus, Guo, Lee, Antin, Perez, Buchanan-King, & Hankey, 2016; Xing, Lv, Wang, Wang, Ai, Cao & Wang, 2019).

Currently, mostly passive safety systems such as airbags and seat-belts have played a large role in the protection of the driver and passengers. Instead of minimizing the injuries after the accident, now many efforts have been devoted to de development of safer and more intelligent systems to prevent accidents from happening (Jarašūnienė & Jakubauskas, 2007; Xing et al., 2019). But these systems usually make decisions without taking driver intended maneuvering into consideration.

Not only driving itself is a complex task, the driver often performs in-vehicle secondary tasks, also called '*multi-tasking*'. The execution of the primary tasks in combination with the in-vehicle secondary task consume extreme visual, cognitive and action resources. This poses a serious threat to drive performance and safety (Lansdown, Brook-Carter & Kersloot, 2004; Chen et al., 2019; Mehler, Reimer, Coughlin & Dusek, 2009). The dual-task paradigm is often used to study the phenomenon of multi-tasking. Multiple studies observe a strong effect of dual-task interference in driving (Cooper, Vladisavljevic, Strayer & Martin, 2008; Strayer, Cooper, Turrill, Coleman, & Hopman, 2017: Broeker, Haeger, Bock, Kretschmann, Ewolds, Künzell & Raab, 2020; Klauer, Guo, Simons-Morton, Ouimet, Lee & Dingus, 2014).

Although the dual-task paradigm is widely used, the effects of primarily cognitively loading tasks on driving performance are not well understood yet (Engström, Markkula, Victor & Merat, 2017). Before we dive into the effects of cognitive loading tasks, it is necessary to give a definition of cognitive load. This can be defined as the ratio between the capacities of the information processing system needed to correctly perform the task and the amount of available attentional resources at any given time (de Waard, 1996; O'Donnell & Eggemeier, 1986). One theory that explains how the ratio is achieved is the cognitive control hypothesis (Engström et al., 2017). This hypothesis states that cognitive load selectively impairs driving subtasks that rely on cognitive control but leaves automatic performance unaffected. Cognitive control is defined as a broad concept for executive cognitive functions such as working memory and attention.

Besides the different secondary tasks that are widely studied, there is a variety of driving forms studied. A specific form of driving is lane changing. A lane change is in this paper defined as a driver maneuver that moves a vehicle from one lane to another where both lanes have the same direction of travel (Fitch, Lee, Klauer, Hankey, Sudweeks & Dingus, 2009). Thus it is a deliberate and substantial shift in the lateral position (Tijerina, Garrott, Stoltzfus & Parmer, 2005). This form is not yet completely understood and multitasking seems to have a negative impact. The paper of van der Heiden, Janssen, Donker and Merkx (2019) studied the phenomenon of lane changing while performing a secondary task. They measured how much time and distance is needed for a full lane change in a naturalistic driving simulator when drivers are placed under different levels of cognitive distraction. The drivers received a warning sign to initiate the lane change. The results showed that some participants took too much time to perform the lane change. In addition, the paper demonstrated that the type of distraction affected the initial reaction time. Van der Heiden et al. (2019) found also differences in mental workload for the dual-task conditions. However the paper did not explain what the relationship is between the perceived workload, reaction time and distance. Besides this, it would be interesting to see if certain aspects of mental workload are more or less responsible for an increased cognitive distraction. This would help develop better human information processing models (Rubio, Diaz, Martin, & Puente, 2004). Furthermore, the paper did not explain which aspect of the steering can account for these effects. The steering process itself is a complex cognitive operation, and it is interesting to understand which aspects of steering are compromised by cognitive load, resulting in the slower reaction times.

This paper is build up twofold. On the one hand it builds upon the article of van der Heiden et al. (2019) and aims to further investigate which component of mental workload is responsible for the decrease in drive performance when there is a higher level of cognitive distraction. In addition, this paper explores the relationship between reaction time, mental workload and distance that is needed to perform a lane change. On the other hand this paper dives deeper into the steering wheel behavior during the lane change. The reaction times that were measured in the paper of van der Heiden et al. (2019) are a combination of several aspects. The underlying cognitive processes which can vary for each reaction time will be studied in this second part. To study cognitive performance, cognitive architectures are widely used to model human behavior (Salvucci, 2001; Salvucci, Boer & Liu, 2001). A cognitive architecture is a general framework for specifying computational behavioral models of human cognitive performance. To observe and predict specific steering wheel behavior during lane change, the two point visual control model of steering of Salvucci & Gray (2004) is used. With this model it is possible to measure and quantify these steering wheel behavior effects.

Due to the fact that there is not yet an explanation for the effects found in the article of van der Heiden et al. (2019), this study has multiple goals by analyzing the existing data in two different ways. Firstly, we focus on individual differences and we will perform two analyses. The first analysis will answer which component of mental workload can be held responsible for the differences in cognitive distraction and therefore driver performance. Secondly, another analysis will be done to find out what the relationship is between the variables reaction time, distance and mental workload. Thirdly, we will take a different view and focus on the underlying cognitive processes. In this second part we will perform a third analysis to look at the steering wheel data in more detail to answer the question which aspect of steering can account for the effects already found.

It is hypothesized that the ratings for workload will differ for the subscales and that mental workload is a predictor for reaction time and distance. Furthermore we hypothesize that the condition in which people drive affects the abruptness of steering wheel behavior.

2. Method

This section summarizes the original study, by giving an overview of its settings, methodology, and results (van der Heiden et al., 2019). We will focus on those aspects of the study that are of relevance for our current purposes. For a full methodological overview we refer to the original study.

2.1 Participants and context

The original study involved twenty-four participants (9 women; 15 men) ranging in age from 28 to 70 years (M = 46.5 years, SD = 12.4 years). The sample matched the distributions of highway drivers in the Netherlands on age, gender and yearly driving distance as found in a population-based study (CBS, 2013).

2.2 Material and stimuli

Participants sat on an adjustable fixed chair in front of a Logitech G27 racing wheel and a 2900 monitor. There were single and dual-task conditions in which the lane change performance was studied. In the dual-task conditions, participants were asked to steer a simulated car while performing an audio task at the same time. We will explain these individual tasks next.

2.2.1 The driving task

The driving task was developed in a modified version of OpenDS 2.5. The driving task was to stay in the middle lane (3.5 m wide) of a straight three-lane highway. A simulated navigation system was shown at the bottom right of the screen. When a lane closure was imminent, the interface showed which lanes were closed (red crosses), following symbols that are used on Dutch highways to indicate closed lanes. Participants were instructed to change to the open lane once they noticed the alert. The open lane was either to the left or the right of the central lane. The visual cue for lane closure was only shown on the interface, not on the road. The car drove at a constant speed of either 80 km/h or 130 km/h.

Within a block of trials, participants changed lanes 6 times (half left, half right). The upcoming lane closure was signaled at one of six locations (225, 275, 325, 375, 425 or 475 m after trial start), which were balanced over trials and subjects. At the end of each trial the car was automatically reset to the middle lane to begin a new trial, and the navigation screen was cleared (showing only a green background).

2.2.2 The audio task

In the dual drive task, an audio task created distraction. There were two audio conditions. In both the audio conditions (i.e., *repeat* and *generate*), participants heard a stream of words, presented at a steady pace of 1 word every 4 s. In the repeat condition, participants simply had to repeat the word they heard. In the generate condition, participants had to respond with a new word that started with the last letter of the word they heard.

2.2.3 Subjective mental workload

The subjective workload that participants experienced during the tasks, was measured with the raw TLX (Task Load Index). This questionnaire consists of six dimensions. These are the mental, physical, temporal, performance, effort and frustration dimension. Table 1 shows the definitions of TLX dimensions (Hart & Staveland, 1988). The scores of these dimensions are combined in one index. Directly after each condition, the participants completed the TLX questionnaire to assess the workload experienced during the task (Hart & Staveland, 1988; Rubio et al., 2004).

Table 1. Rating scale definitions from the TLX.

Title	Description
Mental demand	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical demand	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal demand	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?

Frustration level	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

2.3 Design

For the first analysis, a 2 (Driving speed: 80 km/h, 130 km/h) x 3 (Audio task: No audio, Repeat, Generate) within-subjects repeated measures (ANOVA) design was used to replicate the original study and analyze each subscale of the TLX. Conditions were blocked by speed level. Half the participants started with 80 km/h, the other with 130 km/h. Within each speed condition, participants completed all audio conditions in a specific order. All participants started with the no audio condition (this will be referred to as condition ‘D’). This was followed by the repeat and generate condition (this will be referred to as ‘DH’ and ‘DG’ respectively), of which the order was counterbalanced. The order of audio conditions that was used for the first speed level, was also used for the second speed level. Holm-Bonferroni-corrected post hoc tests were applied in the case of pairwise comparisons.

For the second research question we used a linear mixed effects model (Baayen, Davidson & Bates, 2008). This was done to investigate if the mental workload score could be a predictor for the reaction time and the distance. The dependent variable for the first model was reaction time and the independent variables were mental workload, speed and condition. The model included mental workload, speed (80 and 130 km/h) and condition (D, DH, DG) and a mental workload x speed x condition interaction for fixed effects.

Another linear mixed effects model was constructed to see if mental workload could predict the average distance that was needed for a lane change. The dependent variable for this model was the average distance and the independent variables were mental workload, speed and condition. The model included mental workload, speed and condition and a mental workload x speed x condition interaction for fixed effects.

In both models a random intercept for subject was included to account for within-subject correlations.

Analysis software

Statistics were done using R 3.6.3 (R Core Team, 2020). We used the aov() function which is built into R to analyze the repeated measures ANOVA. We used the lmer program of the lmerTest package for estimating fixed effects and variance/covariance component parameters of the LMM. The full reproducible code is available in Supplementary Materials.

2.4. Procedure

The experiment started with an explanation and the informed consent was signed. Before the experiment started, the participant was able to practice the trials. Participants then performed six experimental blocks. After each block, participants filled out the TLX questionnaire.
The total procedure took approximately 70 min.

2.5. Measurements

To analyze the research question, four facets of behavior were measured:

1. *Initial reaction time* was defined as the time it took before the first steering movement exceeding 1 degree was made (T1 in Fig. 1) relative to the onset of the in-car warning (T0). The initial reaction time is a proxy of how long it takes drivers to initiate a lane change after first stimulus onset. For each condition, we calculated drivers' mean initial reaction time. As the simulator logged steering reaction time since presentation of the visual stimulus, some of these steering movements are not in response to the visual presentation (but due to ongoing steering movements). To compensate for this, we removed reaction times that were logged as being faster than 500 ms.
2. *Lane change distance* was defined as the distance that was traveled between when the in-car warning showed (T0) and the moment at which the car was fully in the target lane (T2). This is shown in figure 1.The criterion for considering the car to be fully in the target lane was that the full body of the car passed the center of the lane markings. Note that the simulator only logged the timestamp of the lane change (i.e., T2), not the distance. The distance was calculated based on the known constant speed and the measured lane change time.
3. *Subjective workload* was measured using the TLX questionnaire, where participants scored their workload on a scale from 1 to 20 on various subscales (Hart & Staveland, 1988). We analyzed the average score of the sum of the TLX, but also each subscale.

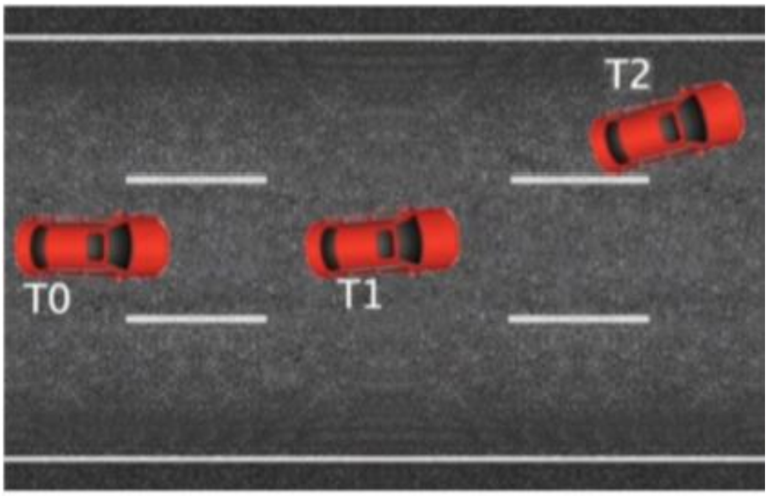


Figure 1. Measurements overview showing the critical timestamps in the experiment. At T0 the visual warning is first presented. At T1, the participant makes the first (bigger) steering action. At T2 the car is fully in the target lane (van der Heiden et al., 2019).

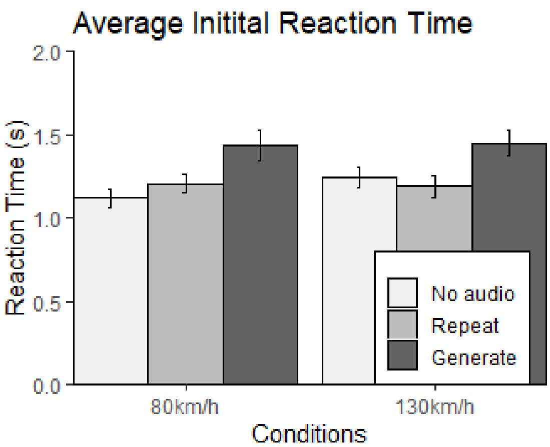
3. Results of novel analysis one and two

3.1 Initial Reaction Time (T1-T0)

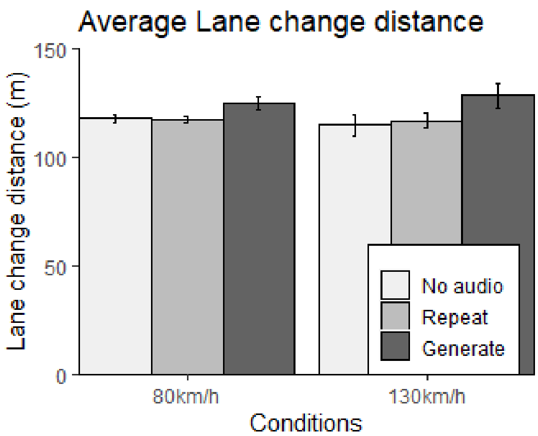
Fig 2A presents the initial reaction time in all conditions (i.e. time interval T1-T0).

There was a significant effect of condition on reaction time, $F(2,46) = 6.85$, $p < .01$, $\eta_p^2 = .23$.

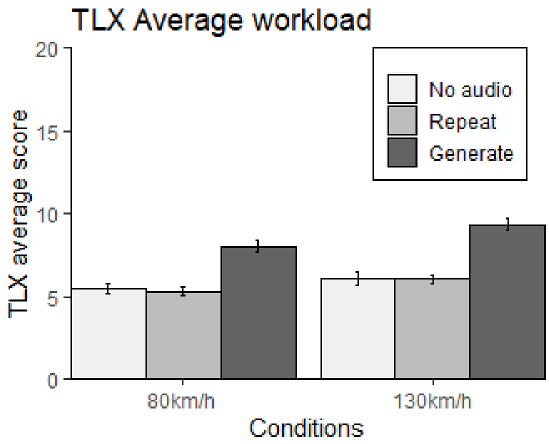
A post hoc test confirmed that reaction times were longer in the *generate* condition ($M = 1.44$ s, $SD =$ .51 s), compared to the *repeat* ($M = 1.20$ s, $SD = 0.32$ s, $p = .01$) and the *no-audio* condition ($M = 1.18$, $SD = 0.34$, $p = .005$). There was no significant effect between the no-audio and repeat condition ($p >$ .1). The initial reaction time at 80 km/h ($M = 1.25$ s, $SD = 0.45$ s) was not significantly different from that at 130 km/h ($M = 1.29$, $SD = 0.39$), $F(1,23) = .41$, $p > .1$. Furthermore, there was no significant interaction between speed and audio condition, $F(2,46) = .97$, $p > .1$.



A



B



C

***Figure 2*. (A)** Bar graph of the average initial reaction time (T1). The graph shows that participants react slower to the visual warning signal in the generate condition compared to the no audio and repeat condition. **(B)** Bar graph of total average lane change distance (at T2) in different conditions. Participants made the lane change in shorter distance when they drove 80 km/h compared to the 130km/h. **(C)** Bar graph of the average TLX score perceived by the participants. The participants experienced more mental load when there was more cognitive distraction.

3.2. Lane change distance (distance covered during T0 until T2)

Figure 2B shows the average lane change distance. There was a significant effect of speed on average lane change distance $F(1,23) = 80.96$, $p < 0.001$, $\eta_p^2 = .78$. Overall, the lane change distance was significantly shorter at 80 km/h (M = 101 m, SD = 17 m) than at 130 km/h (M = 140 m, SD = 31 m). There was also a significant effect of audio condition, $F(2,46) = 3.43$, $p = .04$, $\eta_p^2 = .13$. Although the mean lane change distance was largest in the generate condition, a post hoc test did not find any significant difference between the generate (M = 127 m, SD = 35 m), repeat, (M = 117 m, SD = 29 m) and the no audio condition (M = 116 m, SD = 30 m) (all p >.1). Furthermore there was no significant interaction between speed and audio condition, $F(2,46) = .686$, $p > .1$.

3.3 Average perceived mental workload

Figure 2C shows the perceived mental load indicated on the raw TLX questionnaire. There was a significant difference between the conditions on perceived mental load, $F(2,46) = 41.48$, $p < 0.001$, $\eta_p^2 = .64$. A post hoc test showed that perceived load was significantly higher in the generate (M= 8.7, SD = 3.7) compared to the repeat condition (M = 5.7, SD = 3.2, $p < 0.001$). Perceived mental load was also significantly higher in the *generate* condition compared to the *no-audio* condition (M = 5.8, SD = 3.6; $p < .001$). But there was no significant difference between *repeat* and *no-audio* (p >.1).
There was also a significant effect of driving speed on perceived mental load, $F(1,23) = 8.15$, $p = 0.008$, $\eta_p^2 = .26$. Mental load while driving at 130 km/h was significantly higher (M = 7.13, SD = 3.96) compared to mental load at 80 km/h (M = 6.26, SD = 3.53). There was no significant interaction effect between speed and audio condition, $F(2,46) = 1.31$, $p > .1$.

3.3.1 TLX subscale mental workload

Figure 3A shows the perceived mental load for the TLX_mental subscore indicated on the raw TLX questionnaire. This will be referred to as mental workload. There was a significant difference between the three audio conditions on perceived mental score, $F(2,46) = 37.71$, $p < 0.001$, $\eta_p^2 = .62$. A post hoc test showed that perceived load was significantly higher in the generate ($M = 9.96$, $SD = 4.49$) compared to the repeat condition ($M = 6.42$, $SD = 3.93$; $p < 0.001$). The mental workload was also significantly higher in the generate condition compared to the no-audio condition ($M = 6.4$ $SD = 4.57$; $p < .001$). But there was no significant difference between repeat and no-audio ($p >.1$).
There was also a significant effect of driving speed on mental workload, $F(1,23) = 5.19$, $p = 0.03$, $\eta_p^2 = .18$. Mental load while driving at 130 km/h was significantly higher ($M = 8.08$, $SD = 4.89$) compared to mental load at 80 km/h ($M = 7.1$, $SD = 4.32$). However, there was no significant interaction effect between speed and audio condition, $F(2,46) = 0.72$, $p > .1$.

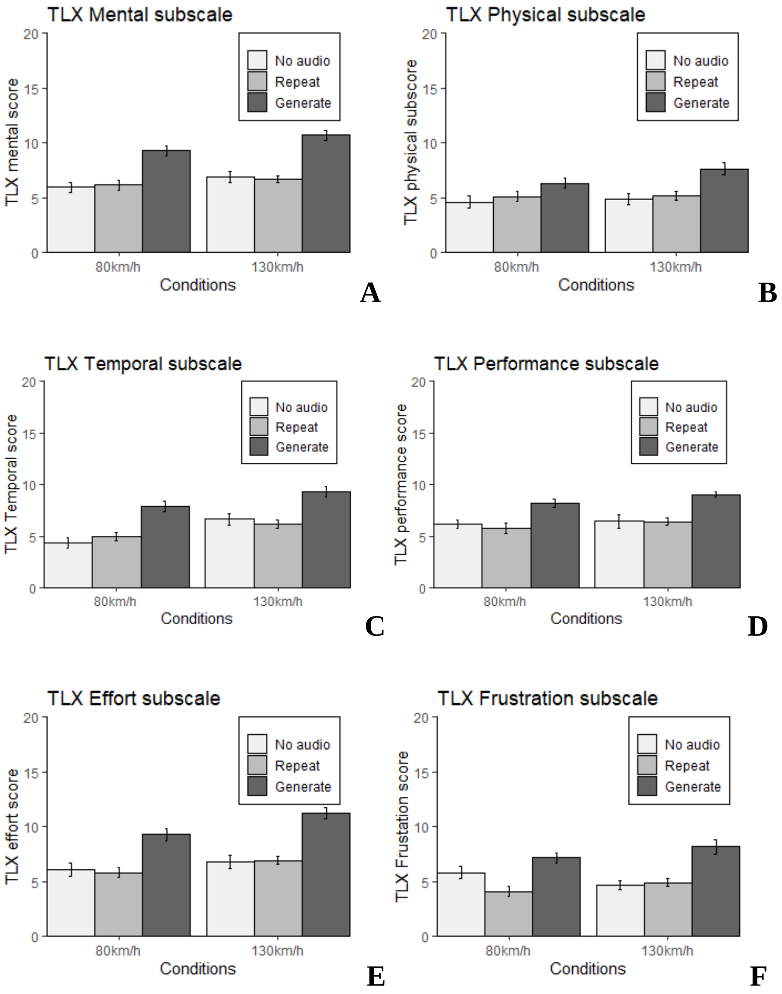


***Figure 3*. A)** Bar graph of the subscale mental workload for each condition. Drivers experienced more mental workload in the generate condition compared to the repeat and no audio condition. And more

in the 130 km/h conditions compared to the 80 km/h conditions. **B)** Bar graph of the subscale physical workload. The bar graph shows that there was a significant effect of condition on physical workload. **(C)** Bar graph of the subscale temporal workload. The bar graph shows drivers experiences significant more temporal workload in the generate condition compared to no audio and repeat and more temporal workload in the 130 km/h conditions than 80 km/h. **(D)** Bar graph of the subscale performance workload for each condition. There was an effect of conditions on performance workload. **(E)** Bar graph of the subscale effort workload for each condition. There was a significant effect of condition and speed on effort workload. **(F)** Bar graph of the subscale frustration workload for each condition. There was a significant effect of condition on frustration workload and a significant interaction effect between condition and speed on frustration workload.

3.3.2 TLX subscale physical workload

Figure 3B shows the perceived mental load for the TLX_physical score indicated on the raw TLX questionnaire. This will be referred to as physical workload. There was a significant effect of condition on physical workload, $F(2,46) = 10.35$, $p < 0.001$, $\eta_p^2 = .31$. A post hoc test showed that physical workload was significantly higher in the generate ($M = 6.96$, $SD = 4.35$) compared to the repeat condition ($M = 5.13$, $SD = 3.37$; $p = 0.06$). The physical workload was also significantly higher in the generate condition compared to the no-audio condition ($M = 4.73$ $SD = 3.71$; $p = 0.02$). But there was no significant difference between repeat and no-audio ($p > .1$).

There was no significant effect of driving speed on physical workload ($p > .1$) There was also no significant interaction effect between speed and audio condition, $F(2,46) = 2.14$, $p > .1$.

3.3.3 TLX subscale temporal workload

Figure 3C shows the perceived mental load for the TLX_temporal score indicated on the raw TLX questionnaire. This will be referred to as temporal workload. There was a significant difference between the three audio conditions on temporal workload, $F(2,46) = 27.18$, $p < 0.001$, $\eta_p^2 = .54$. A post hoc test showed that temporal workload was significantly higher in the generate ($M = 8.60$, $SD = 4.$) compared to the repeat condition ($M = 5.56$, SD = 3.5; $p < .001$). The temporal workload was also significantly higher in the generate condition compared to the no-audio condition ($M = 5.48$ $SD = 4$; $p < 0.001$). But there was no significant difference between repeat and no-audio ($p > .1$).

There was a significant effect of driving speed on temporal workload, $F(1,23) = 9.31$, $p < .001$, $\eta_p^2 = .29$. Temporal workload while driving at 130 km/h was significantly higher ($M = 87.36$, $SD = 4.26$) compared to mental load at 80 km/h ($M = 5.74$, $SD = 3.75$). There was no significant interaction effect between speed and audio condition, $F(2,46) = 1.02$, $p > .1$.

3.3.4 TLX subscale performance workload

Figure 3D shows the perceived workload for the TLX_performance score indicated on the raw TLX questionnaire. This will be referred to as performance workload. There was a significant difference between the three audio conditions on performance workload, $F(2,46) = 23.91$, $p < 0.001$, $\eta_p^2 = .51$. A post hoc test showed that performance workload was significantly higher in the generate ($M = 8.60$, $SD = 3.9$) compared to the repeat condition ($M = 6.06$, $SD = 4.08$; $p = 0.008$). The performance workload was also significantly higher in the generate condition compared to the no-audio condition ($M = 6.29$ $SD = 4.2$; $p = 0.02$). But there was no significant difference between repeat and no-audio ($p > .1$).
There was no significant effect of driving speed on performance workload, $F(1,23) = 1.66$, $p > .1$.
There was no significant interaction effect between speed and audio condition, $F(2,46) = 0.267$, $p > .1$.

3.3.5 TLX subscale effort workload

Figure 3E shows the perceived mental load for the TLX_effort score indicated on the raw TLX questionnaire. This will be referred to as effort workload. There was a significant difference between the three audio conditions on effort workload, $F(2,46) = 26.74$, $p < 0.001$, $\eta_p^2 = .54$. A post hoc test showed that effort workload was significantly higher in the generate ($M = 10.23$, $SD = 4.72$) compared to the repeat condition ($M = 6.35$, $SD = 4.29$; $p < 0.001$). The effort workload was also significantly higher in the generate condition compared to the no-audio condition ($M = 6.40$ $SD = 4.61$; $p < .001$). But there was no significant difference between repeat and no-audio ($p > .1$).
There was also a significant effect of driving speed on effort workload, $F(1,23) = 14.64$, $p < 0.001$, $\eta_p^2 = .39$. Mental load while driving at 130 km/h was significantly higher ($M = 8.29$, $SD = 5.07$) compared to mental load at 80 km/h ($M = 7.03$, $SD = 4.60$). There was no significant interaction effect between speed and audio condition, $F(2,46) = .895$ $p > .1$.

3.3.6 TLX subscale frustration workload

Figure 3F shows the perceived mental load for the TLX_frustration score indicated on the raw TLX questionnaire. This will be referred to as frustration workload. There was a significant difference between the three audio conditions on perceived frustration workload, $F(2,46) = 13.55$, $p < 0.001$, $\eta_p^2 = .37$. A post hoc test showed that frustration workload was significantly higher in the generate ($M = 7.65$, $SD = 4.87$ compared to the repeat condition ($M = 4.48$, $SD = 3.59$; $p = 0.0014$). Frustration workload was also significantly higher in the generate condition compared to the no-audio condition ($M = 5.23$ $SD = 4.45$; $p = .022$). But there was no significant difference between repeat and no-audio ($p > .1$). There was no significant effect of driving speed on frustration workload, $F(1,23) = 0.35$, $p > .1$. However, there was a significant interaction effect between speed and audio condition, $F(2,46) = 5.16$ $p < .01$. $\eta_p^2 = .19$.

3.4 Interim discussion
The goal of the first study was to replicate and expand the study of van der Heiden et al. (2019). The re-analysis of the data from van der Heiden et al. (2019) confirmed their results. In their study it was not clear which component of mental workload was responsible for the increase in subjective workload in the dual task conditions. Therefore we analyzed each subscale of the TLX, to see if a specific subscale would show different results. However, the results of each subscale of the TLX showed that there was not a specific subscale responsible for this phenomenon. The drivers experienced for every subscale more workload when there was more cognitive distraction. Now we move on to the second research question.

3.4 Additional second analysis.
Research question 2: What is the relation between reaction time, average distance and mental workload?

3.4.1 Average mental workload as predictor for reaction time
Results from the linear mixed effects model (LMM) showed a significant main effect for the dual task-generate condition ($\hat{\beta}$ = .49, p = 0.0266). All other main effects and interaction effects were not statistically significant. Table 2 summarizes the results of the LMM for reaction time.

Table 2. Parameter estimates, standard error and t values from the linear mixed model with mental workload, speed and condition as the fixed effects and reaction time as dependent variable.

	Estimate	*SE*	*t*
(Intercept)	1.169***	0.150	7.775
Mental workload	-0.009	0.023	-0.408
Generate	0.494*	0.220	2.247
Repeat	0.099	0.201	0.491
130	-0.085	0.192	-0.444
Mental workload x Generate	-0.019	0.029	-0.666
Mental workload x Repeat	-0.002	0.033	-0.057
Mental workload x 130	0.036	0.029	1.247
Generate x 130	-0.232	0.321	-0.722
Repeat x 130	-0.046	0.280	-0.164
Mental workload x Generate x 130	0.004	0.040	0.099
Mental workload x Repeat x 130	-0.015	0.043	-0.357

*** $p < 0.001$; * $p < 0.05$.

3.4.2 Average mental workload as predictor for the distance
For the second model we looked if mental workload is a predictor for the distance that is needed to perform a lane change. Results from the linear mixed effects model showed that speed was a

significant predictor of distance when the participant drove 130 km/h (see Table 3) ($\hat{\beta}$ = 27.11, p < .05). All other main and interaction effects were statistically nonsignificant. Table 3 summarizes the results of the LMM for distance.

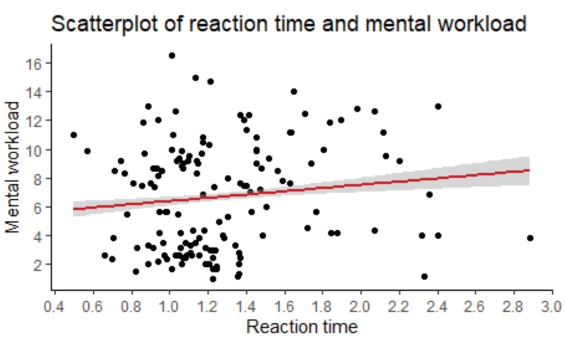
Table 3. Parameter estimates, standard error and t values from the linear mixed model with mental workload, speed and condition as the fixed effects and the distance as dependent variable.

	Estimate	*SE*	*t*
(Intercept)	99.449***	8.866	11.217
Mental workload	-0.199	1.343	-0.148
Generate	-4.529	12.519	-0.362
Repeat	-5.438	11.417	-0.476
130	27.114*	10.916	2.484
Mental workload x Generate	1.513	1.636	0.925
Mental workload x Repeat	0.942	1.852	0.509
Mental workload x 130	1.509	1.631	0.926
Generate x 130	-14.964	18.263	-0.819
Repeat x 130	1.530	15.907	0.096
Mental workload x Generate x 130	1.599	2.251	0.710
Mental workload x Repeat x 130	0.061	2.434	0.025

*** p < 0.001; * p < 0.05.

Since the LMM showed for both models that almost all fixed effects were nonsignificant, a correlation analysis was performed to see if it might be more than just the power causing no effects seen.

A correlation analysis showed (see Figure 4A) that there is a very weak correlation between reaction time and workload (r = .12, p > .1). In addition, the correlation between the distance and mental workload, was also weak (r = .29) but showed more of a trend (see Figure 4B). This suggests that besides the power, there might be no strong relationship in the first place.



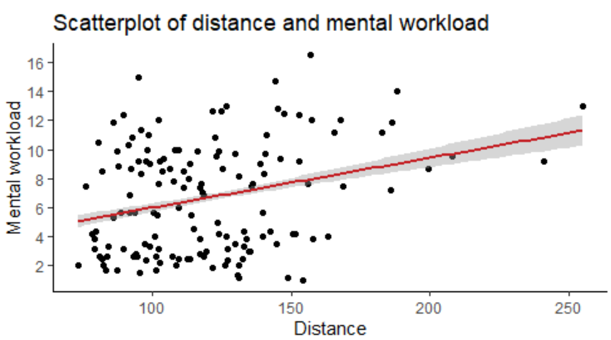


Figure 4. **(A)** Correlation between the reaction time and mental workload. **(B)** Correlation between the distance and mental workload. The red lines indicate regression lines with standard error in grey. For both scatterplots only a weak correlation can be seen.

4. Additional analysis to study the cognitive processes

4.1 Introduction

The goal of the first study is to extend the findings from the first study and to investigate the steering wheel behavior during lane changes in more detail. The variables used in the study of van der Heiden et al. (2019) are the outcome of multiple factors. Thus the underlying cognitive processes which contribute to a T0, T1 or T2 value can vary and in this additional analysis we try to infer these processes. The third research question is therefore: ‘Which aspect of steering accounts for the effects found by van der Heiden et al. (2019)?’.

To observe and predict specific steering wheel behavior during lane change, the two point visual control model of steering (TPVCM) of Salvucci & Gray (2004) was used. This model assumes that when a person drives a car, this driver uses two points to infer position on the road. For this they use a near point and a far point. Information from this far point helps to adjust and compensate for upcoming trajectory of the road. Information of the near point helps to maintain the current lane position of the car. Thus this near point is in nearby distance in front of the car. When people change lanes, the model assumes that both the far and near point change position to the alternate lane. This enables the driver to steer to that lane.

To measure these far and near points, the angle between the direction of the car and the two points is measured. These measurements are quantified in two angles: θ_f *and* θ_n (see figure 5). For every time slot that the location of the car is registered, the model minimizes the change in angle with the near and far point. To maintain a stable position on the road, the model additionally minimizes the angle of the car’s heading with the near point. By showing qualitative comparisons between simulated car movements, and measured car movements, Salvucci and Gray (2004) show that their model can account for a wide range of driving behaviors.

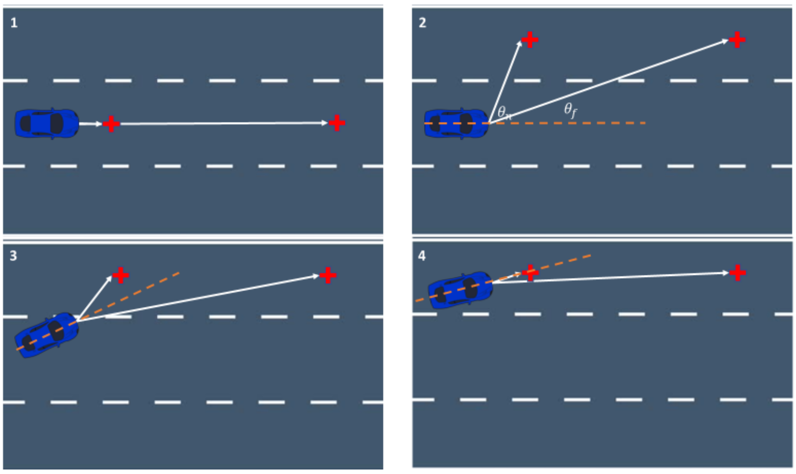


Figure 5. Illustration of a left-ward lane change according to the two-point visual control model of steering. θ_f and θ_n indicate the angle between the car and the far and near points, respectively. Initially, the car drives straight (1). Once the decision to change lanes has been initiated, the far and near points change relative to the car, creating a non-zero angle (2). Consequently, the car steers to the left lane (3), minimizing the angle between the heading of the car and the far and near points (4).

4.2 The Model

Following Salvucci and Gray (2004), we implemented a discrete version of the TPVCM:

$$\Delta\varphi = k_f \Delta\theta_f + k_n \Delta\theta_n + k_I \theta_i \Delta \mathrm{t}$$

In this equation, φ indicates the angle that the car makes relative to its original bearing, θ_f and θ_n indicate the angle between the car and the far and near points, respectively, and $\Delta \mathrm{t}$ represents a time constant of the update cycle. k_f, k_n, and k_I are the contributions of each of these components to the angular change.

In addition to the basic model, we assume a visual processing time that leads to a delayed response to visual events that happen while driving. The implementation of the TPVCM in the ACT-R cognitive architecture introduced a similar mechanism (Salvucci, 2006), but in that model optimization of the delay period to the individual driver is not possible.

The delay period (t_0), together with the parameters governing the steering control (k_f, k_n, and k_I) determine the steering behavior of a driver. Different examples of driving behavior can be seen in figure 2.

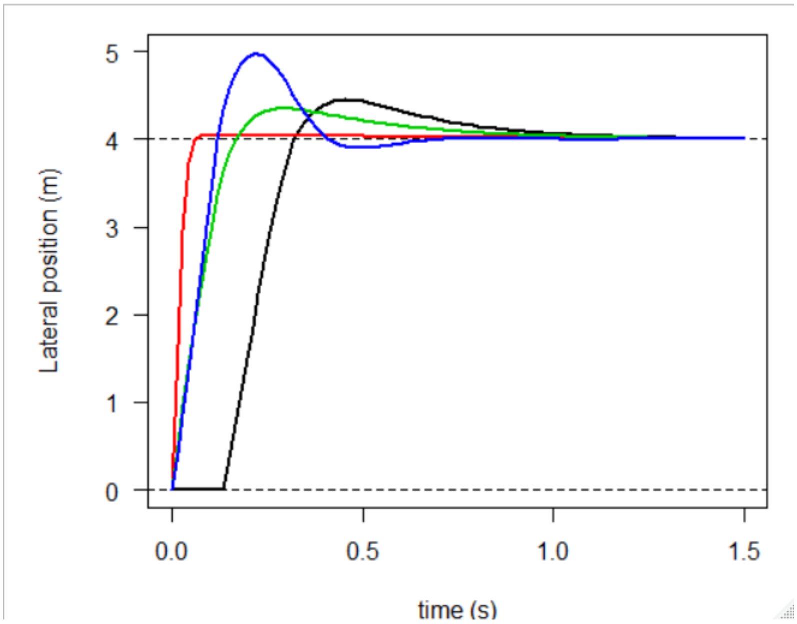


Figure 6. Individual steering profiles predicted by the model. The goal of the model is to steer four meters to the left, beginning at time 0. Black: high t_0; green: high k_f, Red: high k_n, Blue: High k_I.

4.3 Estimating parameters using the two-point model of steering

To reveal individual driver's profiles, we optimize the set of parameters that best describe the steering behavior of every individual. To this end, we predict the lateral deviation of a car under a set of parameters, and minimize the mean squared distance between the observed lateral deviation and the predicted lateral deviation.

The lateral deviation depends on the change in angle and the speed of the car in the following way:

$$\Delta x = v \tan(\varphi) \Delta \mathrm{t}$$

With *v* the speed of the car (in m/s).

Because of the complexity of the parameter space, we applied particle swarm optimization (Clerc, 2010). Preliminary parameter recovery studies with simulated data revealed that the data-generating parameters were recovered with high accuracy.

The parameters were optimized for all time series of all conditions and individuals, excluding trials in which participants either did not change lanes, or started from the incorrect lane (3.4%). All time series were down sampled to 20 Hz, to obtain Δt=50ms, which has been argued is the update time of the human cognitive system (Anderson, 2007; Salvucci & Gray, 2004; Stocco, Lebiere & Anderson, 2010)

4.4 Results

The model fits the data extremely well. Figure 7 (left) shows two example participant's lane deviation during a leftward lane change, which overlaid the model predictions according to the best fitting model parameters. Although the model does not account for brief movements of the car, it captures the overall pattern of the movement. The right panels of Figure 7 illustrate that the model does underestimate the steering angle that the car makes. Nevertheless, the model predicts important features of the steering behavior on which it was not fit. Figure 8 shows that the delay period predicts the initial reaction time T1. The initial reaction time was computed as the time at which the driver makes the first steering motion that exceeds one degree. The delay period seems to be an important component of T1, as it correlates strongly, but is consistently shorter.

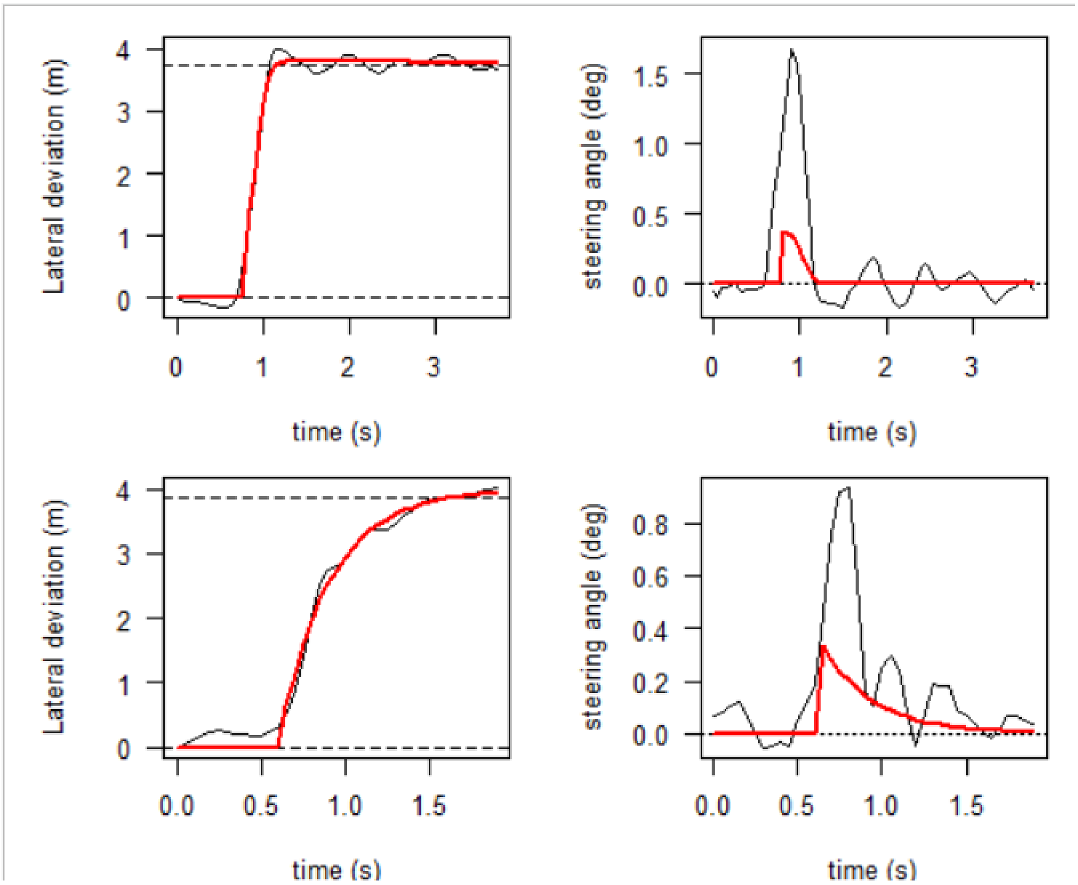


Figure 7. Two example participants with different steering profiles are fit by the model. Left: Observed (black) and predicted lateral deviation from the initial road location. Dashed lines indicate the initial road location and the final road location. Right: Observed (black) and predicted (red) steering angle. Dotted line indicates that steering wheel is in the upright position.

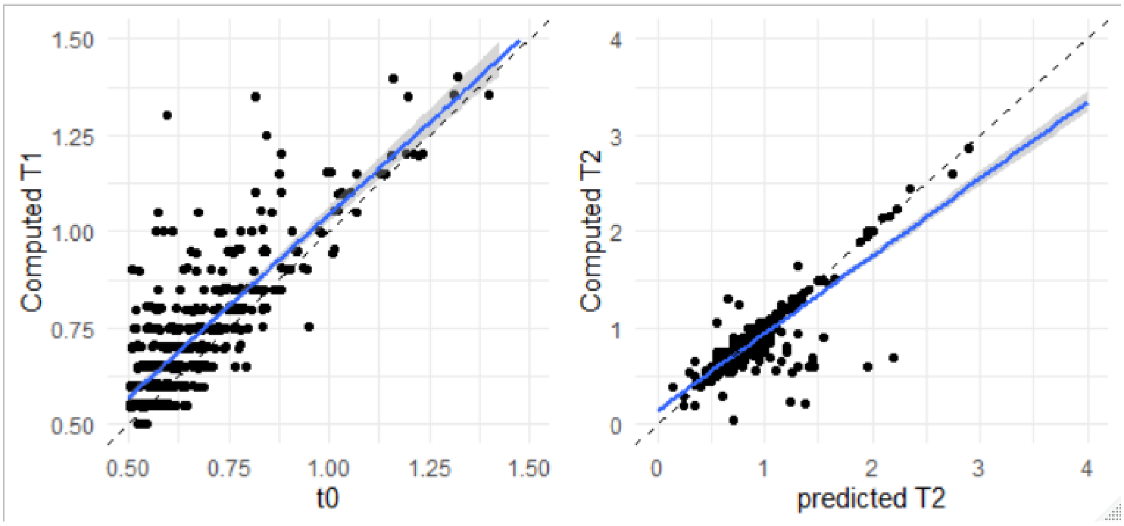


Figure 8. The TPVCM predicts initial (T1) and final (T2) reaction times. A. Correlation between the T1 computed from the data and delay period t_0 from the model. B. Correlation between T2 computed from the data and T2 predicted by the model. The blue lines indicate regression lines; The dashed lines indicate identity lines.

4.5 Materials and stimuli

Design

Now that has been shown that the TPVCM model can be used to estimate and predict the steering behavior, the parameters k_f, k_n, k_I and t_0 will be used for further analysis. For this analysis, a 2 (Driving speed: 80 km/h, 130 km/h) x 3 (Audio task: No audio, Repeat, Generate) within-subjects repeated measures (ANOVA) design was used to analyze all four parameters (k_f, k_n, k_I and t_0) to see if there is an effect of speed and condition for each parameter. A correlation analysis between the parameters is done to see how strong the relationship is between the parameters.

Analysis software

Statistics were done using R 3.6.3 (R Core Team, 2020). We used the 'aov()' function which is built into R to analyze the repeated measures ANOVA. We used the 'ggpairs' function which is built into R to analyze the correlation. The full reproducible code is available in Supplementary Materials.

Measurements

The analysis was done with the four parameters k_f, k_n, k_I and t_0.

- k_f is the parameter for the angular change of θ_f. It contributes to the angular change between the direction of the car and the far point. This indicates that a driver with a relatively high k_f focusses mostly on the far point, and therefore steers more smoothly, taking more time to finish the lane change.

- k_n is the parameter for the angular change of θ_n. It contributes the angular change between the direction of the car and the near point. This parameter adjusts the steering wheel movement on the basis of the change in the angle with the near point relative to the previous time point. A driver with a relatively high k_n minimizes the angle between the car and the near point, resulting in a relatively abrupt steering motion.
- k_I is the parameter for the angular change of θ_i. This parameter adjusts the steering wheel movement relative to the angle with the near point. Thus it enables the driver to steer to the near point. When the k_I is relatively high, the driver tries to maintain the bearing to the near point, there is less focus on the change in angle. Therefore, the steering motion requires overcompensation once the target lane is reached.
- t_0 is the delay period. With this we assume a visual processing time that leads to a delayed response to visual events that happen while driving.

5. Results additional analysis 3

Figure 9A presents the parameter k_f in all conditions. There was a significant effect of speed on k_f $F(1, 19) = 9.98$, $p < .01$. A post hoc test confirmed that the parameter was smaller in the 80 km/h condition ($M = 0.69$, $SD = 0.58$) than the 130 km/h ($M = 1.01$, $SD = 0.540$, $p < .001$). There was no significant effect of condition on k_f, $F(2, 44) = 2.26$, $p > .1$. Furthermore, there was no significant interaction between speed and audio condition, $F(2, 46) = 1.94$, p > .1.

Figure 9B shows parameter k_n in all conditions. There was a significant effect of speed on k_n $F(1, 19) = 101.69$, p < .001. A post hoc test confirmed that the parameter was bigger in the 80 km/h condition ($M = 1.05$, $SD = 0.46$) than the 130 km/h ($M = 0.38$, $SD = 0.31$, $p < .001$). However, there was no significant effect of condition on k_n $F(2, 44) = .60$, $p > .1$ and also no significant interaction effect between speed and condition, $F(2, 46) = 1.58$, $p > .1$.

Figure 9C presents parameter k_I in all conditions. There was a significant effect of speed on k_I $F(1, 19) = 101.65$, $p < .001$. A post hoc test confirmed that the parameter was bigger in the 80 km/h condition ($M = 1.11$, $SD = 0.44$) than the 130 km/h ($M = 0.35$, $SD = 0.32$, $p < .001$). But there was no significant effect of condition on k_I $F(2, 44) = .122$, $p > .1$ and no significant interaction effect between speed and condition $F(2,46) = .621$, $p > .1$.

Figure 9D shows the parameter t_0 in all conditions. There was no significant effect of speed on t_0, F(1,19) = 0.12, p > .1. In addition, there was no significant effect of condition on t_0 F(2,44) = 0.92, p >.1. Finally, there was no significant interaction effect between speed and condition F(2,46) = 2.49, p > .1.

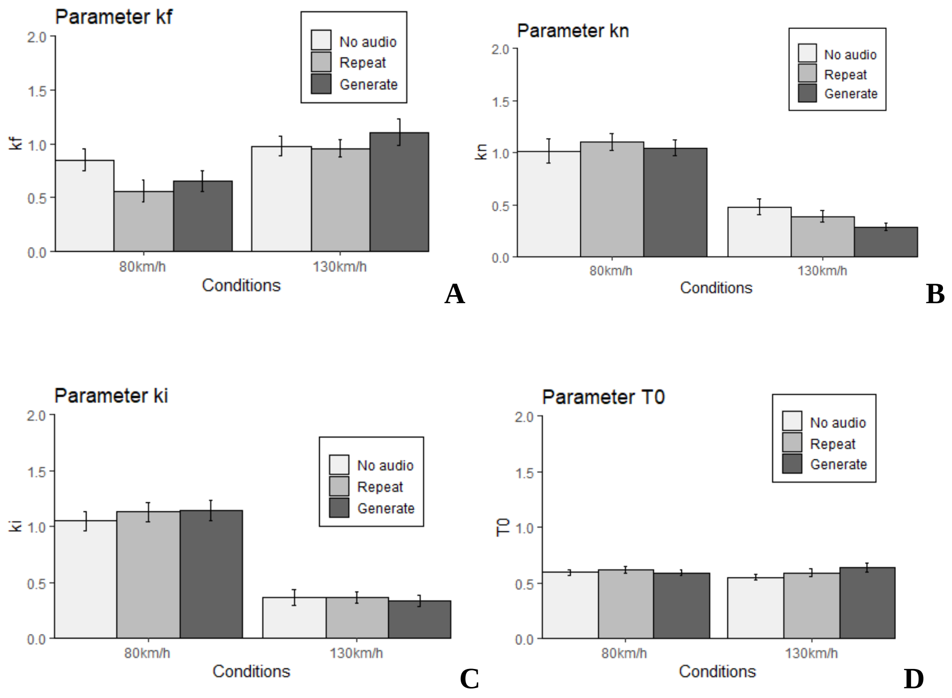
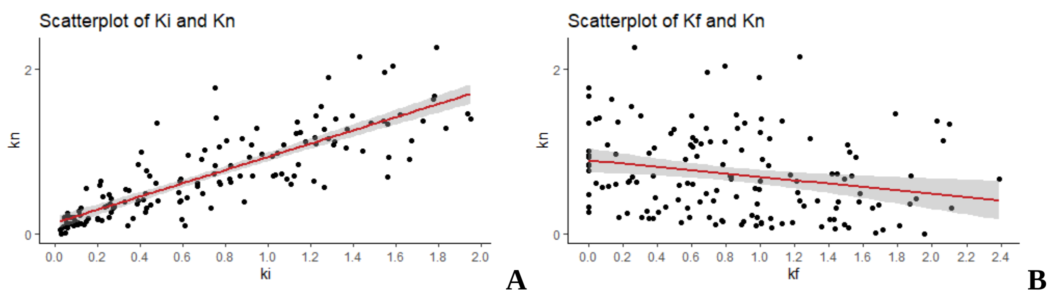


Figure 9**. (A)** Bar graph of the average k_f values. The graph shows that participants have a higher k_f value in the 130 km/h condition. **(B)** Bar graph of the average parameter k_n. The graph shows that there was a significant difference between the speed conditions. **(C)** Bar graph of the average k_I values. The graph shows that participants have a bigger k_I value in the 80km/h condition than in the 130km/h condition. **(D)** Bar graph of the average t_0 values. There was no significant effect found of either speed, condition or the interaction of those.

A correlation analysis showed that there was a significant relationship between k_n and k_I (*r* = .84, *p* < .001). In addition, $k_n$and k_f were significantly correlated, but showed only a weak relationship (*r* = -.23, *p* = .006). t_0 and k_n were not correlated which each other (*r* = .04, *p* > .1). The same results were fount for the correlation analyses between k_f and k_I (*r* = .-.14, *p* >.1), between t_0 and k_I (r = .13, *p* > .1) and between $k_f$and t_0 (*r* = .02, *p* >.1). The visualization of these correlations can be seen in figure 10.



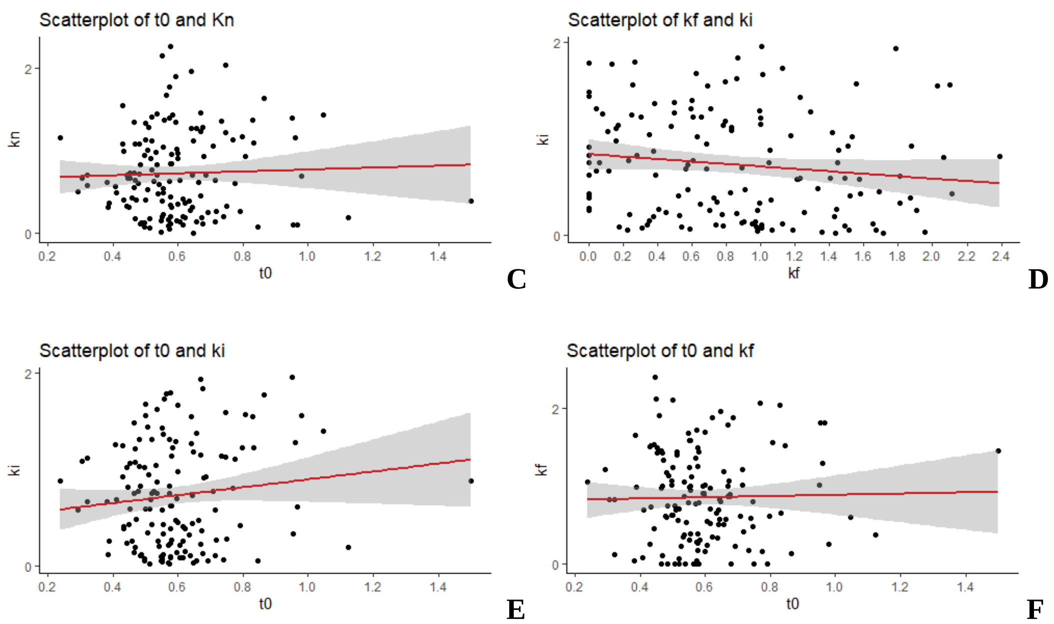


Figure 10. **(A)** The scatterplot shows a strong correlation between k_I en k_n **(B)** Correlation between k_n en k_f **(C)** Correlation between t_0 and k_n **(D)** Correlation between $k_f$en k_I **(E)** Correlation between t_0 en k_I **(F)** Correlation between t_0 en k_f.

6. General discussion

The objective of the current study was twofold because we studied the dataset of van der Heiden et al. (2019) in two different ways. The first part of this paper started with analyses about individual differences in lane change behavior. We checked if a subscale of the TLX questionnaire could be responsible for the increase in cognitive distraction during lane change in different conditions. Furthermore, it was investigated if mental workload was a predictor of reaction time and distance. In the second part of this paper we analyzed the data with a different perspective. We tried to infer the cognitive processes which contribute to the reaction times measured in the first part of the study. This was done with the TPVCM of Salvucci & Gray (2004).

The results of the first part showed that there was not a specific subscale of the TLX questionnaire which was responsible for the decrease in drive performance. This raises the question if the subscales of the TLX questionnaire are distinct enough from each other. The fact that the subscales strongly correlate with each other substantiates this question (Hart & Staveland, 1988). Although the distinction between the subscales remains unclear, the study of Rubio et al. (2004) compared the TLX questionnaire to other subjective measurements and showed that the total TLX scores quite well on sensitivity and validity. This would implicate that only the TLX total scores can be used to measure workload, without specifications of the subscales. Another study of Loeches De La Fuente, Berthelon, Fort, Etienne, De Weser, Ambeck, & Jallais (2019) measured mental workload not only with subjective and behavioral measures, but also with electrophysiological measures. They suggest that an increase of mental workload can have long term effects on performance but shorter effects on electrophysiological measures. To develop better human information processing models and safety systems, both behavioral and physiological parameters should be used.

Besides the interest in mental workload specifically, we also investigated the relationship between reaction time, distance and the mental workload that participants experienced. The linear mixed effects model showed that mental workload was not a predictor of reaction time or distance. Several reasons can explain these results.

First of all, there are different ways in which the analyses could have been done. The limitations of the current dataset made it challenging to meet the criteria for the repeated measures ANOVA, linear regression, correlation or linear mixed effects model. One of the problems was the relatively small sample of participants. Due to this small sample, issues with power arose for almost every analysis.

Secondly, although a lack of power might be part of the problem, the correlation analysis showed that there is not a strong relation between the reaction time, distance and mental workload. Thus, a very real possibility is that in fact there exists no relationship between these factors.

Thirdly, another possible explanation might be the problem of a very controlled task. The driving task, where drivers steered on a lane road without traffic, provided a well-controlled environment in which we looked specifically at the effects of a secondary task on lane change performance. However, it is clearly important to extend this work to more complex domains that would better represent real-world driving situations where more cognitive distraction is available (sight, other cars etc.). This could, in combination with more and different forms of secondary tasks, lead to different results (Pavlidis, Dcosta, Taamneh, Manser, Ferris, Wunderlich & Tsiamyrtzis, 2016).

In the final analysis, a different approach was used to study driving behavior. The two point visual control model of steering (TPVCM) of Salvucci & Gray (2004) was used to answer the third research question and to observe and predict steering wheel behavior. The final analysis showed that the data fitted the model well. An important finding of these results is that the delay period which is used in the model, predicts the initial reaction time (T1) very well. Thus the delay period seems to be very important for T1. Lastly, a novel analysis was done to see if speed and condition had an effect on the parameters of the steering model. The results showed that the values of k_f increased when the participant drove faster, which seems logical since the driver needs more time to steer smoothly to the other lane at higher speed. In addition, for the k_n parameter the opposite effect was seen. This seems again logical, the driver can steer more abruptly at a lower speed. In addition, the $k_I$values increased when the speed was slower. The correlation analysis showed that only k_n and k_I were highly correlated. This seems logical since k_I enables the driver to steer to the near point, and k_n adjusts the steering wheel movement on the basis of the change in the angle with the near point. The results of all these parameter analyses mean that drivers tend to shift from focus from the near point to the far point when they start driving faster. Furthermore drivers steer less abruptly when they drive faster with a focus on the far point. This effect of the speed on steering wheel amplitudes is in line with earlier findings (van Winsum, de Waard, & Brookhuis, 1999; Käppler, 1986). This partly confirms our third hypothesis, namely that speed affects the abruptness of steering wheel behavior.

In conclusion, the goal of current research was to examine lane change behavior in two different ways; first with the focus on individual differences and in the second part with a focus on the underlying cognitive processes. Although we did not find differences in the TLX subscales and can not state that mental workload is a predictor of reaction time or distance, we did find promising results in de second part of the study. We did not only demonstrate that the model from Salvucci & Gray (2004) can be used to observe and predict steering wheel data, we also showed that drivers tend to shift from focus from the near point to the far point when they start driving faster. Furthermore drivers steer less abruptly when they drive faster. This information can be used to design safer roadways and enhance safety systems in cars by incorporating these cognitive processes.

7. Author Contribution Statements

The dataset of van der Heiden et al. (2019) was used for this thesis. The supervisor, L. van Maanen, fitted the data to the model of Salvucci & Gray (2004) and programmed the data in a way that the student Sietske Bootsma could implement the analyses with the parameters. The part which contained the fitting of the dataset (section 4.2 till 4.4) was made by L. van Maanen. Furthermore, the figures 4 till 7 were made by the supervisor.

8. References

Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* Oxford UP.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Broeker, L., Haeger, M., Bock, O., Kretschmann, B., Ewolds, H., Künzell, S., & Raab, M. (2020). How visual information influences dual - task driving and tracking. *Experimental Brain Research*, *238*(3), 675–687. https://doi.org/10.1007/s00221-020-05744-8

Chen, P., Xue, Q., & Jiang, X. (2019). The influence of the secondary tasks on driving performance and mental workload at different speeds. *ACM International Conference Proceeding Series*, 33–39. https://doi.org/10.1145/3352740.3352747

Chong, I., Mirchi, T., Silva, H. I., & Strybel, T. Z. (2014). Auditory and visual Peripheral Detection Tasks and the Lane Change Test with high and low cognitive load. *Proceedings of the Human Factors and Ergonomics Society*, *2014-January* (ISO 26022), 2180–2184. https://doi.org/10.1177/1541931214581458

Clerc, M. (2010). *Particle swarm optimization*. Wiley-ISTE.

Cooper, J. M., Vladisavljevic, I., Strayer, D. L., & Martin, P. T. (2008). Drivers' lane changing behavior while conversing on a cell phone in a variable density simulated highway environment. In *Proceedings of the 87th Annual Meeting of Transportation Research Board*. Washington, DC.

de Waard, D. (1996). The Measurement of Drivers' Mental Workload. Haren: Ph.D. thesis, University of Groningen, Traffic Research Centre.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(10), 2636–2641. https://doi.org/10.1073/pnas.1513271113

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of Cognitive Load on Driving Performance: The Cognitive Control Hypothesis. *Human Factors*, *59*(5), 734–764. https://doi.org/10.1177/0018720817690639

Fitch, G.M., Lee, S.E., Klauer, S., Hankey, J., Sudweeks, J., & Dingus, T. (2009). Analysis of lane-change crashes and near-crashes. *US Department of Transportation, National Highway Traffic Safety Administration*.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. Human Mental Workload, 1, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Jarašūnienė, A., & Jakubauskas, G. (2007). Improvement of road safety using passive and active intelligent vehicle safety systems. *Transport*, *22*(4), 284–289. https://doi.org/10.3846/16484142.2007.9638143

Käppler, W. D. (1986). Using a driving simulator to investigate vehicle handling properties. In: H. P. Willumeit, Human decision making and manual control (pp. 89-99). North-Holland: Elsevier Science Publishers.

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England Journal of Medicine*, *370*(1), 54–59. https://doi.org/10.1056/NEJMsa1204142.

Lansdown, T. C., Brook-Carter, N., & Kersloot, T. (2004). Distraction from multiple in-vehicle secondary tasks: Vehicle performance and mental workload implications. *Ergonomics*, *47*(1), 91–104. https://doi.org/10.1080/00140130310001629775

Loeches De La Fuente, H., Berthelon, C., Fort, A., Etienne, V., De Weser, M., Ambeck, J., & Jallais, C. (2019). Electrophysiological and performance variations following driving events involving an increase in mental workload. *European Transport Research Review*, *11*(1), 1–9. https://doi.org/10.1186/s12544-019-0379-z

Mehler, B., Reimer, B., Coughlin, J. F., & Dusek, J. A. (2009). Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transportation Research Record*, *2138*(1), 6–12. https://doi.org/10.3141/2138-02

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of perception and human performance (Vol. 2 Cognitive processes and performance (pp. 1–49). Oxford: Wiley.

Papantoniou, P., Yannis, G., & Christofa, E. (2019). Which factors lead to driving errors? A structural equation model analysis through a driving simulator experiment. *IATSS Research*, *43*(1), 44–50. https://doi.org/10.1016/j.iatssr.2018.09.003

Pavlidis, I., Dcosta, M., Taamneh, S., Manser, M., Ferris, T., Wunderlich, R., Akleman, E. & Tsiamyrtzis, P. (2016). Dissecting Driver Behaviors Under Cognitive, Emotional, Sensorimotor, and Mixed Stressors. *Scientific Reports*, *6*(1), 1–12. https://doi.org/10.1038/srep25651

Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, *53*(1), 61–86. https://doi.org/10.1111/j.1464-0597.2004.00161.x

Salvucci, D. D. (2001). Predicting the effects of in-car interface use on driver performance: An integrated model approach. *International Journal of Human Computer Studies*, *55*(1), 85–107. https://doi.org/10.1006/ijhc.2001.0472

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, *48*(2), 362–380. https://doi.org/10.1518/001872006777724417

Salvucci, D. D., Boer, E. R., & Liu, A. (2001). Toward an Integrated Model of Driver Behavior in Cognitive Architecture. *Transportation Research Record: Journal of the Transportation Research Board*, *1779*(1), 9–16. https://doi.org/10.3141/1779-02

Salvucci, D. D., & Gray, R. (2004). A two-point visual control model of steering. *Perception*, *33*(10), 1233–1248. https://doi.org/10.1068/p5343

Stocco, A., Lebiere, C., & Anderson, J. R. (2010). Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. *Psychol Rev*, *117*(2), 541–574. https://doi.org/10.1037/a0019077

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J. R., & Hopman, R. J. (2017). The smartphone and the driver's cognitive workload: A comparison of Apple, Google, and Microsoft's intelligent personal assistants. *Canadian Journal of Experimental Psychology*, *71*(2), 93–110. https://doi.org/10.1037/cep0000104

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing Cognitive Distraction in the Automobile. *Human Factors*, *57*(8), 1300–1324. https://doi.org/10.1177/0018720815575149

Tijerina, L., Garrott, W. R., Stoltzfus, D., & Parmer, E. (2005). Eye glance behavior of van and passenger car drivers during lane change decision phase. *Transportation Research Record*, *1937*, 37–43. https://doi.org/10.3141/1937-06

van der Heiden, R. M. A., Janssen, C. P., Donker, S. F., & Merkx, C. L. (2019). Visual in-car warnings: How fast do drivers respond? *Transportation Research Part F: Traffic Psychology and Behaviour*, *65*, 748–759. https://doi.org/10.1016/j.trf.2018.02.024

van Winsum, W., de Waard, D., & Brookhuis, K. A. (1999). Lane change manoeuvres and safety margins. *Transportation Research Part F: Traffic Psychology and Behaviour*, *2*(3), 139–149. https://doi.org/10.1016/s1369-8478(99)00011-x

Xing, Y., Lv, C., Wang, H., Wang, H., Ai, Y., Cao, D., Velenis, E., & Wang, F. Y. (2019). Driver Lane Change Intention Inference for Intelligent Vehicles: Framework, Survey, and Challenges. *IEEE Transactions on Vehicular Technology*, *68*(5), 4377–4390. https://doi.org/10.1109/TVT.2019.2903299

9. Supplementary materials

Research question 1: reaction time

```
###stap1 : inladen van allData file
load("~/Thesis/allData file/allData.Rdata")
View(allData)
###stap 2: RT_change1 aanpassen
RTchange1 <- replace(allData$RT_change1, allData$RT_change1<100, NA) #alle waarden onder 100
        eruit halen en NA maken
#de juiste, staat niet in allData
RTchange1 <- RTchange1/1000 #de juiste, staat niet in allData
###stap 3: nieuw dataframe gemaakt met de nieuwe kolom reactietijden
allDatanew <- cbind(allData, RTchange1)
###stap 4: verwijder allData
###stap 5: eerst gemiddelde RTchang1 berekenen
meanRT <- aggregate(RTchange1 ~ speed + trial + PPN, mean, data=allDatanew)
###stap 6: repeates measurements anova uitvoeren voor RTchange1
anova1 <- (aov(RTchange1 ~ speed*trial + Error(PPN/(speed*trial)), data=meanRT))
summary(anova1)
###stap 7: effectsize berekenen
EtaSq(anova1, type = 1, anova = TRUE)
#posthoc bonferroni
pairwise.t.test(meanRT$RTchange1, meanRT$trial, p.adjust.method = "bonferroni")
#descriptive statistics voor gemiddelde en sd per conditie
aggregate(RTchange1 ~ trial, meanRT, mean )
aggregate(RTchange1 ~ trial, meanRT, sd )
#descriptive statistics per speed
aggregate(RTchange1 ~ speed, meanRT, mean )
aggregate(RTchange1 ~ speed, meanRT, sd )
```

Research question 1: distance

```
###stap1 : inladen van allData file
load("~/Thesis/allData file/allData.Rdata")
View(allData)
###stap2 : variabelen ordenen
RTchange4 <- allData$RT_change4/1000 #de juiste, zelfde maar dan in ms
##berekenen van average lane change distance
#de km/u omrekenen naar meters per seconde
```

```
speednew <- allData$speed/3.6
#afstand in meters per seconde
averagedist <- RTchange4 * speednew
#kolom met average distance toevoegen aan allData
allDatanew <- cbind(allData, averagedist)
#verwijder allData (de oude) en RTchange4 en speednew
#gemiddelde afstand berekenen per conditie
meandist<- aggregate(averagedist ~ trial + speed + PPN, mean, data = allDatanew)
#repeates measurements anova uitvoeren voor averagedist=gemiddelde afstand nodig voor lane
    change
anova2 <- (aov(averagedist ~speed*trial + Error(PPN/(speed*trial)), data=meandist))
summary(anova2)
#effect size
EtaSq(anova2, type = 1, anova = TRUE)
#posthoc bonferroni trial
pairwise.t.test(meandist$averagedist, meandist$trial, p.adjust.method = "bonferroni")
#posthoc bonferroni speed
pairwise.t.test(meandist$averagedist, meandist$speed, p.adjust.method = "bonferroni")
#descriptive statistics per conditie
aggregate(averagedist ~ trial, meandist, mean)
aggregate(averagedist ~ trial, meandist, sd)
#descriptive statistics per speed
aggregate(averagedist ~ speed, meandist, mean)
aggregate(averagedist ~ speed, meandist, sd)
```

Research question 1: mental workload

```
#inladen excel file
Results <-read_excel("~/Thesis/resultsv_subjective/Resultsv_subjective_adapt2.xlsx")
View(Results)
Results = Results[!is.na(Results$`TLX mental`),]
View(Results) #nu zijn alle NA's verwijderd
#gemiddelde maken van de tlx average score
Results$tlxaverage <- Results$`TLX average` # geen spatie meer, zodat je later geen probleem krijgt
meanTLXaverage <- aggregate( tlxaverage ~ speed + Task + Participant, mean, data = Results)
#anova uitvoeren
anova3 <- (aov(tlxaverage ~ speed * Task + Error(Participant/(speed * Task)), data =
    meanTLXaverage))
```

```
summary(anova3)
#effect size
EtaSq(anova3, type = 1, anova = TRUE)
#posthoc bonferroni trial
pairwise.t.test(meanTLXaverage$tlxaverage, meanTLXaverage$Task, p.adjust.method =
    "bonferroni")
#descriptive statistics per conditie
aggregate(tlxaverage ~ Task, meanTLXaverage, mean)
aggregate(tlxaverage ~ Task, meanTLXaverage, sd)
#posthoc bonferroni speed
pairwise.t.test(meanTLXaverage$tlxaverage, meanTLXaverage$speed, p.adjust.method =
    "bonferroni")
#descriptive statistics per speed
aggregate(tlxaverage ~ speed, meanTLXaverage, mean)
aggregate(tlxaverage ~ speed, meanTLXaverage, sd)
```

Research question 1: TLX mental subscale

```
###TLX mental
#gemiddelde maken van de tlx mental scores
Results$tlxmental <- Results$`TLX mental` # geen spatie meer, zodat je later geen probleem krijgt
meanTLXmental <- aggregate( tlxmental ~ speed + Task + Participant, mean, data = Results)
#anova uitvoeren
anova4 <- (aov(tlxmental ~ speed * Task + Error(Participant/(speed * Task)), data =
    meanTLXmental))
summary(anova4)
```

Research question 1: TLX physical subscale

```
Results$tlxphysical <- Results$`TLX physical` # geen spatie meer, zodat je later geen probleem krijgt
meanTLXphysical <- aggregate( tlxphysical ~ speed + Task + Participant, mean, data = Results)

#anova uitvoeren
anova5 <- (aov(tlxphysical ~ speed * Task + Error(Participant/(speed * Task)), data =
    meanTLXphysical))
summary(anova5)
```

Research question 1: TLX temporal subscale

```
Results$tlxtemporal <- Results$`TLX temporal` # geen spatie meer, zodat je later geen probleem
    krijgt
```

```
meanTLXtemporal <- aggregate( tlxtemporal ~ speed + Task + Participant, mean, data = Results)

#anova uitvoeren
anova6 <- (aov(tlxtemporal ~ speed * Task + Error(Participant/(speed * Task)), data =
        meanTLXtemporal))
summary(anova6)
```

Research question 1: TLX effort subscale

```
###TLX effort
Results$tlxeffort <- Results$`TLX effort`# geen spatie meer, zodat je later geen probleem krijgt
meanTLXeffort <- aggregate( tlxeffort ~ speed + Task + Participant, mean, data = Results)

#anova uitvoeren
anova8 <- (aov(tlxeffort ~ speed * Task + Error(Participant/(speed * Task)), data = meanTLXeffort))
summary(anova8)
```

Research question 1: TLX frustration subscale

```
###TLX frustration
#gemiddelde maken van de tlx frustration scores
Results$tlxfrustration <- Results$`TLX frustration` # geen spatie meer, zodat je later geen probleem
        krijgt
meanTLXfrustration <- aggregate( tlxfrustration ~ speed + Task + Participant, mean, data = Results)

#anova uitvoeren
anova9 <- (aov(tlxfrustration ~ speed * Task + Error(Participant/(speed * Task)), data
        =meanTLXfrustration))
summary(anova9)
```

Research question 1: TLX Performance subscale

```
#gemiddelde maken van de tlx performance scores
Results$tlxperformance <- Results$`TLX performance` # geen spatie meer, zodat je later geen
        probleem krijgt
meanTLXperformance <- aggregate( tlxperformance ~ speed + Task + Participant, mean, data =
        Results)

#anova uitvoeren
```

```
anova7 <- (aov(tlxperformance ~ speed * Task + Error(Participant/(speed * Task)), data =
    meanTLXperformance))
summary(anova7)
```

Research question 2: LMM

```
###stap 1: dataset inladen
load("~/Thesis/allData file/allData.Rdata")
###stap 2: de juiste variabelen
#RT_change1
RTchange1 <- replace(allData$RT_change1, allData$RT_change1<100, NA) #alle waarden onder 100
    eruit halen en NA maken
RTchange1 <- RTchange1/1000 #nog delen door 1000, de juiste, staat niet in allData
#RT_change4
RTchange4 <- allData$RT_change4/1000 #de juiste, zelfde maar dan in ms
#nieuw dataframe gemaakt met de nieuwe RTchange1 en RTchange4
allDatanew <- cbind(allData, RTchange1, RTchange4)
#verwijderen allData dataframe
###stap 3: eerst gemiddelde RTchang1 berekenen
meanRT <- aggregate(RTchange1 ~ speed + trial + PPN, mean, data=allDatanew)
###stap 4: gemiddelde average distance berekenen
#stap4a: afstand berekenen
#de km/u omrekenen naar meters per seconde
speednew <- allDatanew$speed /3.6
#afstand in meters per seconde
averagedist <- RTchange4 * speednew
#verwijderen van speednew en RTchange4
#kolom met average distance toevoegen aan allDatanew
allDatanew <- cbind(allDatanew, averagedist)

#stap 4b: gemiddelden berekenen adhv aggregate
#de gemiddelden afstand berekenen per proefpersoon
meanaveragedist <- aggregate(averagedist ~ speed + trial + PPN, mean, data=allDatanew)
###stap 5: dan gemiddelde meanmentalworkload berekenen
library(readxl)
#eerst file laden
Results <-read_excel("~/Thesis/resultsv_subjective/Resultsv_subjective_adapt2.xlsx")
```

```
Results = Results[!is.na(Results$`TLX mental`),] #alle NA's verwijderen
#even anders benoemen
Results$tlxaverage <- Results$`TLX average`
meanmentalworkload <- aggregate(tlxaverage ~ speed + Task + Participant, mean, data=Results)
### lmer met mental workload as predictor for RTchange1
analyse <- lmer(meanRT$RTchange1 ~ meanmentalworkload$tlxaverage * meanRT$trial *
    as.factor(meanRT$speed) + (1|meanRT$PPN))
summary(analyse)
EtaSq(analyse, type = 1, anova = TRUE)
###lmer met mental workload as predictor for average distance
analyse2 <- lmer(meanaveragedist$averagedist ~ meanmentalworkload$tlxaverage *
    meanaveragedist$trial * as.factor(meanaveragedist$speed) + (1| meanaveragedist$PPN))
summary(analyse2)
```

Research question 3;

```
load("~/Thesis/deel 2 stuurmodel/fit3.Rdata")
####Analyse voor parameters####
###---kf----###
#stap 1: gemiddelde score berekenen voor kf
meanparameterkf <- aggregate(kf ~ speed + cond + ID, mean, data = fit)
#stap 2: anova uitvoeren voor kf
anova1 <- (aov(kf ~ speed*cond + Error(ID/(speed*cond)), data=fit))
summary(anova1)
#posthoc
pairwise.t.test(meanparameterkf$kf, meanparameterkf$speed, p.adjust.method = "bonferroni")
aggregate( meanparameterkf$kf ~ speed, meanparameterkf, mean )
aggregate( meanparameterkf$kf ~ speed, meanparameterkf, sd )


###----kn----###
#stap 3: gemiddelde score berekenen voor kn
meanparameterkn <- aggregate(kn ~ speed + cond + ID, mean, data = fit)
#stap 4: anova uitvoeren voor kn
anova2 <- (aov(kn ~ speed*cond + Error(ID/(speed*cond)), data=fit))
summary(anova2)
#posthoc
pairwise.t.test(meanparameterkn$kn, meanparameterkn$speed, p.adjust.method = "bonferroni")
```

```
aggregate( meanparameterkn$kn ~ speed, meanparameterkn, mean )
aggregate( meanparameterkn$kn ~ speed, meanparameterkn, sd )
###----ki----###
#stap 5: gemiddelde score berekenen voor ki
meanparameterki <- aggregate(kI ~ speed + cond + ID, mean, data = fit)
#stap 6: anova uitvoeren voor kI
anova3 <- (aov(kI ~ speed*cond + Error(ID/(speed*cond)), data=fit))
summary(anova3)
#posthoc
pairwise.t.test(meanparameterki$kI, meanparameterki$speed, p.adjust.method = "bonferroni")
aggregate( meanparameterki$kI ~ speed, meanparameterki, mean )
aggregate( meanparameterki$kI ~ speed, meanparameterki, sd )
####----t0----####
#stap 7: gemiddelde score berekenen voor T0
meanparameterT0 <- aggregate(t0 ~ speed + cond + ID, mean, data = fit)
#stap 8: anova uitvoeren voor T0
anova4 <- (aov(t0 ~ speed*cond + Error(ID/(speed*cond)), data=fit))
summary(anova4)
```